DOCUMENT RESUME

ED 128 854 CS 501 484

AUTHOR Louden, Allan; Shellen, Wesley N.
TITLE A Study of Ballot Feedback from Professional and Nonprofessional Debate Judges.
PUB DATE 76
NOTE 12p.

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS *Cocurricular Activities; *Debate; Educational Research; Evaluation Criteria; *Feedback; Higher Education; Nonprofessional Personnel; Professional Personnel; *Public Speaking; *Scoring

ABSTRACT

Variations in scoring and feedback provided by professional and nonprofessional debate judges were studied using 197 ballots from the six preliminary rounds of the Big Sky Intercollegiate Forensic Tournament in 1971. Results indicated no significant difference in the sides (affirmative or negative) awarded the decision by the two groups of judges. However, nonprofessional judges awarded significantly more quality points to both individuals and teams than did professional judges. Nonprofessional judges were also less likely to provide written feedback to the debaters, wrote shorter comments when they did provide feedback, and relied more heavily upon nonsubstantive, evaluative comments. There was no evidence to suggest that the nonprofessional judges were incompetent to judge the debates, but they may have been less secure about their duties. (AA)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED128854

A STUDY OF BALLOT FEEDBACK FROM PROFESSIONAL AND NONPROFESSIONAL DEBATE JUDGES

Allan Louden--Director of Forensics
Northwest Community College, Powell, Wyoming

and

Wesley N. Shellen--Assistant Professor
University of Montana, Missoula, Montana

Evaluation and feedback are important in any educational debate situation. We assume debaters are aided in learning skills through accurate feedback and, in most instances, the critics communicate this feedback only through the use of a written ballot. It is hoped that the information provided by ballots will aid the student in evaluating and thus improving his performance.

As everyone knows, we do not always put a great deal of faith in the accuracy of these comments. It is not at all uncommon for debaters, and for that matter their coaches, to question the credibility of judges, particularly the so-called "unqualified judge."

Aside from the question of whether competitive debate should use the professional or nonprofessional judge, most tournament directors are necessarily in the position of using a large body of nonprofessionals (i.e., non-coaches) as judges.

Research to date has found that the professional and nonprofessional judges differ in several ways. A number of studies have tended to show that while coaches and non-coaches

almost always reach the same decisions, they are not for the same reasons.[1] In general, professional judges found "argument" to be the major factor in determining a winner, while nonprofessional judges found "delivery" and delivery-related areas most influential.[2] Little research, however, has focused on the differences in feedback provided by the written ballot. Specifically, would significant differences be expected between professional and nonprofessional judges in: tendency to award the decision to one side, amount of quality points awarded, team ratings assigned, and amount of comments or feedback on ballots. In other words, what variations in feedback can be expected from professional versus nonprofessional judges?

This study focuses on a comparison of professional and nonprofessional judges in providing feedback by means of the written ballot. Such an analysis has several implications. Since most debaters engage in intercollegiate competition, learning how one will be judged and how to better interpret the ballot feedback will be useful. Also, because decisions on who will advance to elimination rounds of tournaments are often close, the variations of professional and nonprofessional judges may significantly affect it. It could make possible, then, greater fairness in assigning judges.

Method

The data for this study were gathered in the six preliminary rounds at the Big Sky Intercollegiate Forensic Tournament

hosted by the University of Montana, April, 1971. The tournament is an open-invitational and attracts competition from approximately thirty colleges largely in a five-state Northwest region. The debates utilized the Parliamentary style format[3] on the national debate topic, "Resolved: That the federal government should establish a program of wage and price controls." Half of the debates were judged by professional judges and half were judged by nonprofessional judges. All judges used Form C of the American Forensic Association ballot and were given an instruction sheet describing the format of the debate and defining the six criteria contained on the ballot.

The professional judges were judges provided by the guest schools. This group consisted largely of forensic coaches and graduate forensic assistants. The nonprofessional judges were volunteer adults obtained from the surrounding community. This group consisted largely of business and professional people and faculty members from various departments in the university. None of the judges were paid.

The professional judges could generally be categorized as being experienced both in the theory and practice of debate and in judging the national debate topic. The nonprofessional judges lacked one or both of these qualifications.

The judges were arbitrarily assigned to rounds with no conscious attempt to select judges for particular debates. Random assignment in its purest sense could not be employed, however, since it was impossible to control the judging times

chosen by the volunteer nonprofessional judges.

The total number of ballots obtained for analysis was 98 ballots (35 senior division; 63 junior division) from the professional judges and 99 ballots (30 senior division; 69 junior division) from the nonprofessional judges. If a judge failed to fill out some scale on the ballot, the data were omitted in the analysis of results concerning that particular scale.

Results

The results are presented as answers to particular research questions asked by the investigators:

1. Do professional and nonprofessional judges tend to award the decision to one side more often than the other?

Table I presents the number of ballots awarded to the affirmative or negative sides by the professional and nonprofessional judges.

Table I
Frequency of Wins Awarded to Sides by Judges

	Professional Judges	Nonprofessional Judges
Affirmative win	58	45
Negative win	41	52

The data from Table I were analyzed using the Chi-Square Test corrected for continuity[4] yielding a nonsignificant $\chi^2 = 2.45$. The data were further subdivided to allow analysis of only the ballots from the senior division ($\chi^2 = 2.31$) and junior division ($\chi^2 = 2.14$), both of which were nonsignificant. This analysis revealed that there were no significant differences in the sides (affirmative or negative) awarded the decision by professional versus nonprofessional judges.

2. Do professional and nonprofessional judges differ in the number of quality points awarded to the individual speakers in the debates?

Most coaches and debaters are familiar with the 30-point rating boxes provided on the AFA ballot, form C, to rate each speaker on a scale from one to five points ("poor" to "superior") on each of six criteria (analysis, reasoning, evidence, organization, refutation, and delivery). Since these quality points or speaker points are often used to break ties to determine finalists among teams with identical win-loss records, it is important to know whether the type of judge employed in the debate biases these ratings. Table II presents the mean speaker ratings by the professional and nonprofessional judges and t-tests of the differences between the means for each division.

The results of the data in Table II showed that nonprofessional judges tended to award significantly higher quality points to the individual speakers in both divisions than did the professional judges.

Table II
Mean Quality Points Awarded to Speakers in Each Division

Source of Speaker Ratings	Professional Judges	Nonprofessional Judges	t
Sr. Div. only n=260	21.0	23.0	3.96*
Jr. Div. only n=528	20.2	21.6	4.21*
Both Divisions (combined)	20.5	22.1	5.49*

*$p < .05$

3. Do professional and nonprofessional judges differ in the quality rating assigned to the teams?

The AFA ballot, form C, also provides a simple five-point scale (ranging from "poor" to "superior") for rating the quality of each of the two teams in the debate. In the tournament, 393 such team ratings were reported on the ballots. Professional judges awarded a mean team rating of 3.41 and nonprofessional judges awarded a mean team rating of 3.66. Although the mean ratings of the two types of judges differed by only a quarter of a point, this difference was statistically significant ($t=3.34$, $p < .05$), indicating that the nonprofessional judges awarded significantly higher quality ratings to teams than the professional judges.

4. Do professional and nonprofessional judges differ in the amount of written feedback made on the ballot?

The AFA ballot, form C, provides spaces for written

comments regarding each of the four speakers in the debate. Other than the actual win-loss judgment, these written critiques have the potential to be the most specific and helpful part of the feedback from the judge to the debaters available on the ballot. Several analyses were made of the ballots from the Big Sky tournament to determine the extent to which the professional and nonprofessional judges utilized this opportunity to communicate in writing to the debaters.

Perhaps nothing is more frustrating to a debater or coach than to read a decision of win or lose on a ballot but to receive no explanation or comments from the judge to indicate why he made such a decision or what he liked or disliked about the debaters' performances. A simple analysis was performed to determine whether professional and nonprofessional judges differed in the tendency to hand in blank, no comment ballots. None of the 99 ballots provided by the professional judges were completely blank but 22 of the 98 ballots provided by nonprofessional judges were completely blank in the spaces provided for comments. Table III shows the number of ballots coming from professional and nonprofessional judges which were lacking in feedback for at least one debater versus those in which comments were provided for all four debaters.

The statistical analysis of the data in Table III showed a highly significant tendency in which professional judges more often provided written feedback to all four debaters than did the nonprofessional judges ($\chi^2 = 25.49$; $p < .001$).

Table III
Frequency of Ballots from Professional Versus Nonprofessional Judges Containing Feedback Versus No Feedback for Debaters

Description of Feedback	Professional Judges	Nonprofessional Judges
Comments for all four debaters	94	64
Blank ballot or no comment for at least one debater	5	34

A random sample of 25 ballots from the professional judges and 25 ballots from the nonprofessional judges was taken to allow a word count of the written comments to further describe the amount of feedback provided by the judges. The mean number of words per ballot written by professional judges was 148 words (ranging from 46 words to 541 words), while the nonprofessional judges wrote a mean of 57 words per ballot (ranging from zero to 154 words). It appeared that the professional judges provided nearly three times as much written feedback as the nonprofessional judges. Indeed, the lengthiest set of comments from the nonprofessional judges (154 words) was only slightly higher than the average of the professional judges (148 words).

Of course, the presence or absence of comments is only a crude measure of the quality of feedback provided by the judges. However, in reading the comments in the sample of 50 ballots used in the word count analysis, the authors agreed

that the professional judges tended more often to provide substantive comments and critiques on the cases and debate techniques employed by the debaters, whereas the nonsubstantive, evaluative comments such as "good debate" and "you have an excellent voice" tended more often to come from the nonprofessional judges. Clearly, the professional judges provided a greater amount of written feedback to the debaters in contrast to the nonprofessional judges.

Discussion

From the results of this study, a general picture of the comparison and contrast of professional and nonprofessional judges emerged. The composite of the professional judge appeared to be an individual who had no particular preference for the side of the question being argued; who tended to be somewhat stingy in rating the individuals and teams; and who tended to provide a substantial amount of written feedback on the ballot. The composite nonprofessional judge appeared to be an individual who also had no particular preference for the side of the question being argued; who tended to be somewhat generous in rating the individual speakers and teams; and who tended to provide little or no written feedback on the ballot.

In tournaments where both professional and nonprofessional judges are used, an attempt should be made to give each team a proportionate number of each type of judge to offset the differences in rating behavior if such ratings are ever to be

used as tie-breakers for finalist selections. If such a procedure is not possible, the first criterion for breaking ties should be based on a system such as median ratings rather than mean or total points, to offset the bias of a generous or stingy judge.

There was no evidence in this study to suggest that the nonprofessional judges were incompetent to judge the debates. The stereotyped image of the conservative townsperson, unwilling to give a fair hearing to the radical affirmative proposal, did not appear in the data. At most, the results suggested that the nonprofessional judges may have been somewhat insecure about their duties, thus tending to avoid writing comments on the ballots which would reveal the bases of their decisions. The practice of using nonprofessional judges to judge "unimportant" debates should probably not be carried to extremes, therefore, since the debaters in such debates are usually those who could profit most from the extensive and substantive critique of a trained professional judge.

Notes

1. Donald Klopf, Diana Evans, and Sister Mary Linus De Lozier, "Comparative Studies of Students, Laymen and Faculty Members as Judges of Speech Contests," Speech Teacher, XIV (November, 1965), 314-318; also Frederick Williams, Sally Ann Webb, and Ruth Anne Clark, "Dimensions of Evaluation in High School Debate," Central States Speech Journal, XVII (February, 1966), 18-19.

2. Williams, Clark, and Barbara Sundene Wood, "Studies in the Dimensionality of Debate Evaluation," JAFA, III (September, 1966), 95-103; also James Roever and Kim Giffin, "A Study of the Use of Judging Criteria in Tournament Debate," AFA Register, VIII (Winter, 1960), 13.

3. Parliamentary style debate is similar to Oxford style but uses seven-minute constructive speeches; followed by a fourteen-minute Parliamentary discussion; and is concluded by one four-minute summary-rebuttal per team. The Parliamentary discussion is used for refutation, rebuttal, or cross-examination and no speaker may hold the floor for more than one minute at a time.

4. Sidney Siegel, Nonparametric Statistics for the Behavioral Sciences (New York: McGraw-Hill Book Company, Inc., 1956), 107.